

基于高光谱技术与机器学习的新疆红枣品种鉴别

刘立新^{1,2*}, 何迪¹, 李梦珠¹, 刘星³, 屈军乐⁴

¹西安电子科技大学物理与光电工程学院, 陕西 西安 710071;

²中国科学院西安光学精密机械研究所瞬态光学与光子技术国家重点实验室, 陕西 西安 710119;

³深圳技术大学中德智能制造学院, 广东 深圳 518118;

⁴深圳大学物理与光电工程学院, 光电子器件与系统教育部/广东省重点实验室, 广东 深圳 518060

摘要 为实现对红枣品种的判别, 利用高光谱技术并结合机器学习算法对金丝大枣、骏枣和滩枣这三个品种的新疆红枣进行研究。首先, 分别利用多元散射校正(MSC)、标准正态变量变换(SNV)、一阶导(1-Der)和 Savitzky-Golay(SG)平滑等数据预处理方法对原始光谱进行预处理, 研究了预处理方法对建模的影响; 然后, 利用光谱-理化值共生距离法(SPHY)将样本集划分为校正集和预测集, 基于线性判别分析(LDA)、K-最近邻分类(KNN)和支持向量机(SVM)算法对预处理后的全波段光谱建立红枣品种鉴别模型, 结果显示, 在多种预处理方法中, 1-Der 的处理效果最好; 然后, 结合主成分分析(PCA)、连续投影算法(SPA)和竞争性自适应重加权采样(CARS)等特征提取方法对全波段光谱进行特征波段的提取, 再基于特征波段建立红枣品种鉴别模型, 结果发现, 在几种特征提取方法中, 基于 CARS 所提特征波段建立的模型可以获得最高的鉴别准确率; 最后, 以 SVM 模型为例对模型运行时间进行了比较, 结果发现, 基于特征波段所建模型的运行时间远短于基于全波段所建模型的运行时间。

关键词 光谱学; 高光谱技术; 机器学习; 品种鉴别; 数据预处理; 特征波段提取

中图分类号 O433.4

文献标志码 A

doi: 10.3788/CJL202047.1111002

Identification of Xinjiang Jujube Varieties Based on Hyperspectral Technique and Machine Learning

Liu Lixin^{1,2*}, He Di¹, Li Mengzhu¹, Liu Xing³, Qu Junle⁴

¹School of Physics and Optoelectronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China;

²State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, Shaanxi 710119, China;

³Sino-German College of Intelligent Manufacturing, Shenzhen Technology University, Shenzhen, Guangdong 518118, China;

⁴College of Physics and Optoelectronic Engineering, Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, Shenzhen University, Shenzhen, Guangdong 518060, China

Abstract To identify different Xinjiang jujube varieties, a hyperspectral technique and machine learning algorithms were employed to obtain and analyze the spectral data of Jinsī-jujube, Jun-jujube, and Tan-jujube. First, the original spectra were preprocessed using various data preprocessing methods, including multiplicative scatter correction (MSC), standard normal variate transformation (SNV), first-derivative (1-Der), and Savitzky-Golay (SG) smoothing. The effects of the preprocessing methods on modeling were investigated. Then, the samples were divided into calibration and prediction sets using sample set partitioning methods based on joint X-Y distance (SPHY). The jujube variety identification models were established based on linear discriminant analysis (LDA), K-nearest neighbor (KNN), and support vector machine (SVM) algorithms using the preprocessed full-band spectra. The results demonstrate that 1-Der outperformed other preprocessing methods mentioned above. Next, the

收稿日期: 2020-04-16; 修回日期: 2020-05-29; 录用日期: 2020-06-23

基金项目: 国家自然科学基金(61378091)、高等学校学科创新引智计划、深圳大学光电子器件与系统教育部/广东省重点实验室开放基金(GD201711)、瞬态光学与光子技术国家重点实验室开放基金(SKLST201804)

* E-mail: lxliu@xidian.edu.cn

characteristic bands were extracted from the full-band spectra using principal component analysis (PCA), successive projections algorithm (SPA), and competitive adaptive reweighted sampling (CARS). Then, the jujube variety identification models were established based on the characteristic bands. The CARS-based models achieved the highest accuracy in the models established based on several characteristic band extraction methods. Finally, taking the SVM model as an example, the model runtime was compared. The time required by the SVM model based on the characteristic bands was much shorter than the time required by the model based on the full-band spectra.

Key words spectroscopy; hyperspectral technique; machine learning; variety identification; data preprocessing; characteristic band extraction

OCIS codes 300.6170; 300.6550; 300.6340

1 引 言

新疆红枣为新疆特有的地理标志产品之一,被称为“黄金寿枣”,具有补气养血的功效。民间有“一日食仁枣,百岁不显老”之说,高度赞扬了红枣的食补功效。不同产地的红枣,由于土壤、水质和气候等影响因素不同,在形状、色泽和品质等方面具有一定差异;不同产地的红枣,其生物学特性不同,且其口感、所含微量元素和营养价值等也不尽相同,而这些差异无法通过肉眼直接进行辨别。目前,市场上红枣产地和品种造假的现象日趋严重,一方面影响了枣农的经济利益,另一方面也严重侵害了消费者的权益。因此,对红枣的品质、产地和品种等进行有效鉴别具有极重要的意义。

传统的农产品检测方法通常会对待测产品造成一定程度的破坏,或者需要消耗较长的检测时间,而且很难实现批量检测。因此,亟需一种新的检测方法对红枣品种进行快速无损检测。高光谱技术兴起于20世纪80年代初期,是新一代光电无损检测技术。它通过测量样品的光谱信息,在一定程度上反映被测样品内部物理结构和化学成分的差异,具有较高的分辨率,可获得物体连续、窄波段的光谱信息。此外,高光谱设备操作简单,分析成本较低,且不会对被检测对象造成破坏。该技术早期主要应用于空间遥感与测量^[1-2],近年来才逐步应用于农产品检测^[3-5]、生物医学诊断和指导^[6-8]、植被和水资源调控^[9-10]等领域,并且都取得了一定的研究成果。

国内外学者利用高光谱技术对苹果、柑橘和桃子等水果的品质进行了定性判别和定量检测^[11-18]。本文基于高光谱技术和机器学习对金丝大枣、骏枣、滩枣三种新疆红枣进行鉴别。本文首先分别采用多元散射校正(MSC)、标准正态变量变换(SNV)、一阶导(1-Der)和 Savitzky-Golay (SG)平滑预处理方法对高光谱数据进行预处理,然后分别利用线性判别分析(LDA)、K-最近邻分类算法(KNN)和支持向

量机(SVM)三种定性判别算法建立红枣品种的鉴别模型,研究不同的预处理方法对模型性能的影响,找出对应模型效果最好的预处理方法。此外,本文结合主成分分析(PCA)、连续投影算法(SPA)和竞争性自适应重加权采样(CARS)三种特征提取方法提取全波段光谱的特征波段,然后基于特征波段建立红枣品种鉴别模型,并将其与基于全波段所建模型进行比较分析,找到了高效、可靠的红枣品种鉴别模型。另外,本文还对基于全波段和特征波段所建模型的运行时间进行了比较。

2 实验材料与仪器

2.1 实验材料

实验所用新疆红枣样本分别为当年产金丝大枣309颗、骏枣284颗和滩枣315颗,果实大小均匀,形状规则,无损伤和病虫害。首先,清洗所有样本表面的污渍,阴干后编号;然后利用光谱-理化值共生距离(SPHY)样本划分方法按照3:1的比例将不同品种的红枣样本划分为校正集和预测集。

2.2 实验仪器

高光谱系统的工作原理如图1所示。本文实验中用到的主要仪器为 Maya2000 Pro 型光谱仪及 ISP-REF 反射式积分球。Maya2000 Pro 型光谱仪的光谱分辨率为 0.5 nm,工作波长为 165 ~ 1100 nm。积分球光源为卤钨灯,它可提供均匀的表面照明,用于测量固体物质的反射率或光谱源的发射。

在进行光谱采集时,首先利用标准白板进行校正,然后将每个样本的赤道部位放置于 ISP-REF 积分球的样本采集口处,保持样本位置不变直至光谱采集结束。利用光谱仪在每一个样本的不同位置处共采集3次光谱数据,取平均值作为该样本的最终光谱值。由于光谱在 165 ~ 400 nm 和 1000 ~ 1100 nm 范围内受噪声的影响较大,故最终选用 400 ~ 1000 nm 波段对光谱数据进行处理和分析。

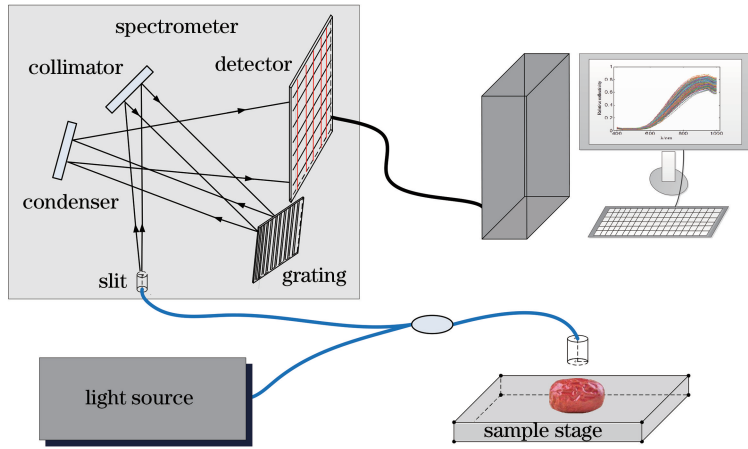


图 1 高光谱系统原理图

Fig. 1 Schematic of hyperspectral system

3 光谱数据的处理

3.1 光谱数据的预处理

光谱数据中通常包含着由相机或仪器产生的随机噪声和光谱变化。对光谱进行预处理可以减少或消除数据中与自身性质无关的信息,降低模型的复杂性,提高数据和模型的可解释性(鲁棒性和准确性)。因此,对光谱数据进行预处理是多变量分析之前必不可少的操作。仪器产生的随机噪声会因样品的物理性质不同而不同,故本实验分别利用 MSC、SNV、1-Der 和 SG 平滑光谱预处理方法对反射光谱数据进行预处理,并对结果进行分析比较,以选取效果最优的预处理方法,减小噪声的干扰,提高信噪比。

3.2 特征波段的提取

典型的高光谱数据由许多波段组成,这些波段通常会表现出高度的相关性并含有大量冗余信息及噪声。特征波段提取方法是从原有的光谱数据中提取与样品化学物质相关的波段,可在一定程度上降低信息的共线性和数据冗余,减少输入数据变量的个数。本文分别采用 PCA、SPA 和 CARS 这三种特征提取方法对特征波段进行提取,并对基于所选特征波段建立的模型进行分析比较,选出最优的特征波段提取方法。

3.3 数据建模

本文利用 SPXY 划分方法将样本集划分为校正集和预测集之后,基于校正集采用 LDA、KNN 和 SVM 三种算法分别建立红枣品种的鉴别模型;然后利用所建立的鉴别模型对预测集样本进行鉴别,并根据鉴别准确率对模型性能进行评价和比较,以选出最优的数据建模方法。模型鉴别准确率的计算公

式为

$$Q = \frac{J}{J_1} \times 100\%, \quad (1)$$

式中: Q 为模型的鉴别准确率,%; J_1 为某品种样本的实际数量; J 为该品种鉴别正确的数量。

4 结果与讨论

4.1 光谱数据的预处理

在光谱数据采集过程中,对每一颗红枣样本在其不同位置共采集 3 次光谱数据,取平均值作为该样本的光谱值;对每种样本的所有光谱曲线求平均值,分别得到金丝大枣、骏枣和滩枣样本的平均光谱曲线,如图 2(a) 所示。所有红枣样本的原始光谱(未经光谱预处理方法处理)曲线如图 2(b) 所示,经过 MSC、SNV、1-Der 和 SG 平滑方法预处理后的光谱如图 2(c)~(f) 所示。在 400~1000 nm 波长范围内,红枣样本的光谱曲线形状很相似,无法直观地从光谱曲线上对品种进行区分,因此还需通过建模对其进行鉴别。

4.2 基于光谱预处理的建模分析

分别采用多种数据预处理方法对原始光谱进行预处理,并基于原始光谱和预处理光谱对经 SPXY 划分的样本校正集和预测集建立 LDA、KNN(选择 k 的值为 4,距离度量采用的是欧几里得距离)和 SVM(选择径向基核函数作为核函数)红枣品种鉴别模型。各模型的鉴别结果如表 1~3 所示。

由表 1~3 可以看出:基于原始光谱所建立的 LDA、KNN 和 SVM 红枣品种鉴别模型的总准确率分别为 65.35%、82.89% 和 95.61%;在几种预处理方法中,基于 1-Der 所建立的红枣品种鉴别模型的总准确率均为最高,分别是 76.32%、100% 和

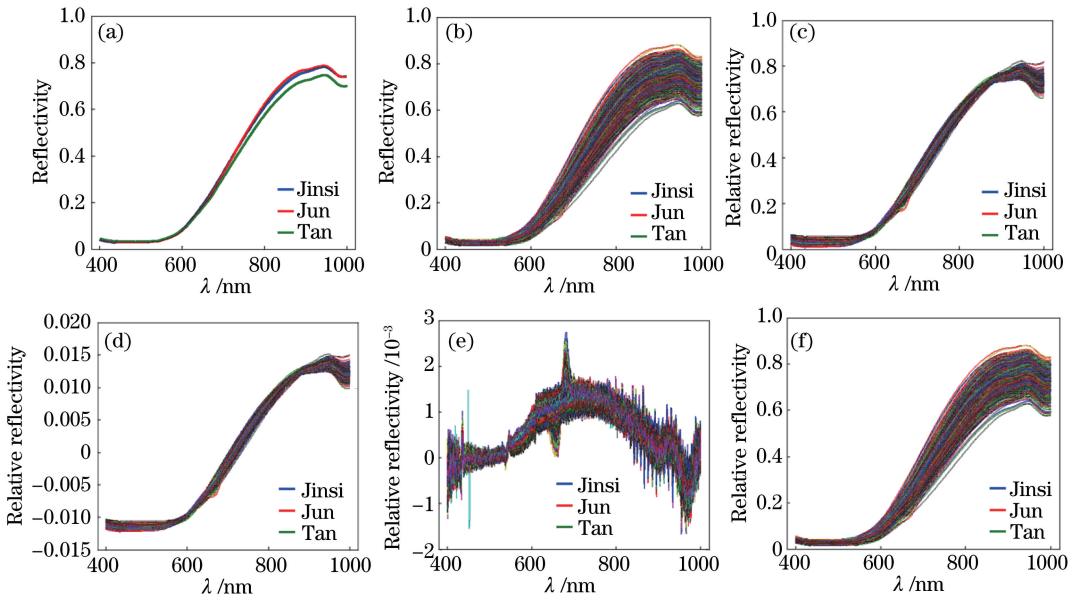


图 2 红枣样本预处理前后的光谱分布。(a)平均光谱;(b)原始光谱;(c) MSC 预处理后的光谱;
(d) SNV 预处理后的光谱;(e) 1-Der 预处理后的光谱;(f) SG 平滑预处理后的光谱

Fig. 2 Spectral profiles of Jujube samples before and after preprocess. (a) Average spectra; (b) original spectra; (c) preprocessed spectra by MSC; (d) preprocessed spectra by SNV; (e) preprocessed spectra by 1-Der; (f) preprocessed spectra by SG smoothing

表 1 基于不同预处理方法所建 LDA 模型的鉴别结果

Table 1 Identification results of LDA model with different pretreatment methods

Pretreatment method	Number of misjudgment			Overall accuracy / %
	Jinsi jujube	Jun jujube	Tan jujube	
Original	37	22	20	65.35
SG smoothing	37	22	20	65.35
MSC	29	26	35	60.53
SNV	29	25	35	60.96
1-Der	29	14	11	76.32

表 2 基于不同预处理方法所建 KNN 模型的鉴别结果

Table 2 Identification results of KNN model with different pretreatment methods

Pretreatment method	Number of misjudgment			Overall accuracy / %
	Jinsi jujube	Jun jujube	Tan jujube	
Original	18	16	5	82.89
SG Smoothing	18	16	5	82.89
MSC	14	12	12	83.33
SNV	14	12	12	83.33
1-Der	0	0	0	100

表 3 基于不同预处理方法所建 SVM 模型的鉴别结果

Table 3 Identification results of SVM model with different pretreatment methods

Pretreatment method	Number of misjudgment			Overall accuracy / %
	Jinsi jujube	Jun jujube	Tan jujube	
Original	5	4	1	95.61
SG Smoothing	3	3	2	96.49
MSC	2	4	2	96.49
SNV	2	4	2	96.49
1-Der	0	0	0	100

100%，优于基于原始光谱以及其他几种光谱预处理方法所建立的模型。这说明光谱预处理方法的选取对于所建模型的准确度有着重要的影响。在后续的红枣品种鉴别分析中，本文选择 1-Der 预处理方法对光谱进行预处理，以提高准确率。

4.3 特征波段的提取

4.3.1 PCA 提取特征波段

利用 PCA 进行主成分分析，得到了金丝大枣、骏枣和滩枣这三个品种红枣样本的前三个主成分

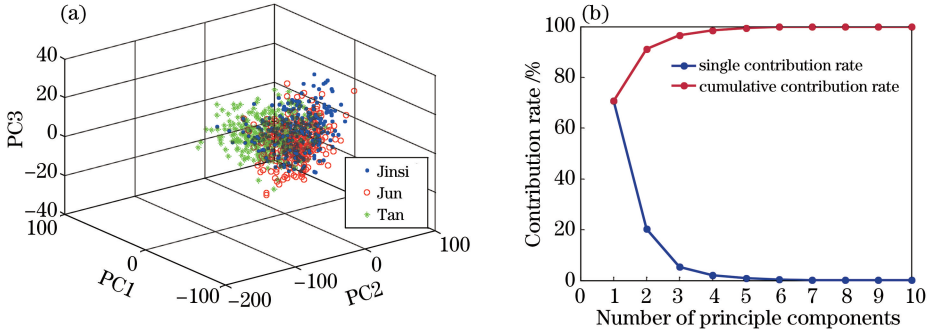


图 3 利用 PCA 法提取特征波段。(a)前 3 个主成分的得分图；(b)前 10 个主成分的方差贡献率
Fig. 3 Extracting characteristic bands by PCA. (a) Scores of the first three principal components; (b) variance contribution rate of the first ten principal components

4.3.2 SPA 提取特征波段

采用 SPA 在红枣样本的高光谱数据中选取特征波段，设置数据不缩放，选取的波段数最小设为 2，最大设为 30。图 4 所示为根据选取的特征变量的个数计算得到的均方根误差 (RMSE)，采用 SPA 算法选出的最佳特征波段数为图中“□”对应的横坐标。即当选取的波段数为 13 时，RMSE 最小，因此基于 SPA 最终可生成只有 13 个特征波段的光谱数据。

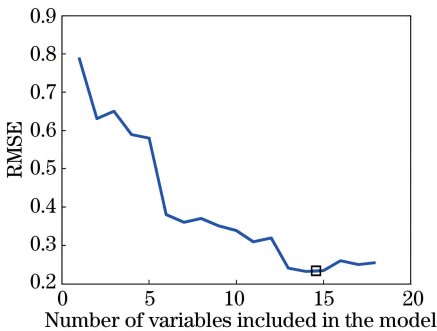


图 4 根据选取的特征变量的个数计算得到的均方根误差
Fig. 4 Root-mean-square error calculated according to the number of selected characteristic variables

4.3.3 CARS 提取特征波段

在 CARS 特征提取算法中设置蒙特卡罗采样的次数为 50。CARS 依据交互验证均方根误差 (RMSECV) 最小的原则进行特征波段的提取，该算

(PC1~PC3) 的空间分布散点图，如图 3(a) 所示。可见，相邻类别的样本空间分布存在着一定数量的交叉，很难直接对不同品种的红枣样本进行鉴别。图 3(b) 为前 10 个主成分的方差贡献率，其中前三个主成分的累积方差贡献率达到了 96% 以上，可以在一定程度上概括原始光谱的主要特征信息。为保证尽可能少地损失光谱的有效信息，本文最终选取前 10 个主成分作为特征变量用于后续的红枣品种鉴别。

法的运算过程如图 5 所示。由图 5(a) 可以看出：在变量提取过程中，被挑选波段的数量随采样次数增加而逐渐下降；在采样前期，波段变量数下降的速度较快，而在后期，变量数下降的速度减慢。图 5(b) 所示为 RMSECV 随采样次数的变化，在采样次数为 14 时 RMSECV 取得最小值 (图 5(c) 中“*”垂线对应的位置)。图 5(c) 为变量回归系数的变化路径，可见，基于第 14 次采样获得的变量子集的均方根误差最小。因此，本文选择该次采样所选择的变量组作为 CARS 最优的特征波段变量 (共 275 个)。

4.4 基于全波段/特征波段的建模分析

对经过 1-Der 预处理后的全波段光谱 (FS)，以及利用多种特征提取方法提取的特征波段的光谱数据，采用 SPXY 划分样本校正集和预测集，分别建立基于 LDA、KNN 和 SVM 的红枣品种鉴别模型。鉴别结果分别列于表 4~6 中。

由表 4~6 可以看出：基于全波段建立的 LDA、KNN 和 SVM 模型的总准确率分别为 76.32%、100% 和 100%；在几种特征提取方法中，基于 CARS 提取的特征变量所建立的 LDA、KNN 和 SVM 模型的总准确率均为最高，分别为 85.53%、98.68% 和 98.25%。可见，在红枣品种鉴别时采用基于 CARS 提取特征波段建模可获得良好的效果。

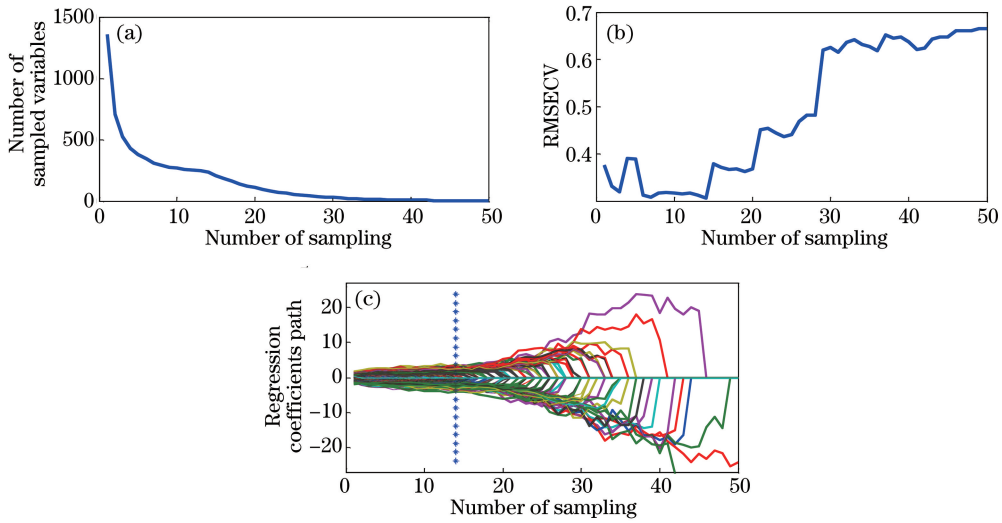


图 5 利用 CARS 提取特征波段。(a)变量数随采样次数的变化曲线；
(b) RMSECV 随采样次数的变化曲线；(c)变量回归系数的变化路径

Fig. 5 Extracting characteristic bands by CARS. (a) Variation curve of the number of variables with the number of sampling; (b) variation curve of RMSECV with the number of sampling; (c) variation path of variable regression coefficient

表 4 基于不同特征波段提取方法所建 LDA 模型的鉴别结果

Table 4 Identification results of LDA model with different characteristic bands extraction methods

Characteristic bands extraction method	Number of misjudgment			Overall accuracy / %
	Jinsi jujube	Jun jujube	Tan jujube	
FS	29	14	11	76.32
PCA	13	20	19	77.19
SPA	18	16	10	80.70
CARS	15	12	6	85.53

表 5 基于不同特征波段提取方法所建 KNN 模型的鉴别结果

Table 5 Identification results of KNN model with different characteristic bands extraction methods

Characteristic bands extraction method	Number of misjudgment			Overall accuracy / %
	Jinsi jujube	Jun jujube	Tan jujube	
FS	0	0	0	100
PCA	10	10	9	87.28
SPA	9	11	6	88.60
CARS	0	3	0	98.68

表 6 基于不同特征波段提取方法所建 SVM 模型的鉴别结果

Table 6 Identification results of SVM model with different characteristic bands extraction methods

Characteristic bands extraction method	Number of misjudgment			Overall accuracy / %
	Jinsi jujube	Jun jujube	Tan jujube	
FS	0	0	0	100
PCA	8	10	7	89.04
SPA	4	5	3	94.74
CARS	2	2	0	98.25

4.5 模型运行时间的比较

衡量所建模型的性能除了考虑准确率之外,模型运行时间也是一个非常重要的参量。本文基于 Windows 7(64 位操作系统), Intel(R) Pentium(R) CPU G645 @ 2.90 GHz 处理器, 利用 MATLAB

2014a 软件对运行时间进行了分析。这里以所建立的 SVM 红枣品种鉴别模型为例, 对基于不同特征提取方法选择的特征波段作为 SVM 模型的输入变量, 得到了基于 SVM 的红枣品种鉴别模型的运行时间, 如表 7 所示。

表 7 基于不同特征波段提取方法的 SVM 模型的准确率和运行时间

Table 7 Accuracy and runtime of SVM model based on different characteristic bands extraction methods

Characteristic bands extraction method	Number of characteristic bands	Accuracy /%	Runtime /s
FS	1356	100	1.497
PCA	10	89.04	0.026
SPA	13	94.74	0.032
CARS	275	98.25	0.167

由表 7 可以看出:基于全波段建模的 SVM 模型的运算时间为 1.497 s,远长于基于特征波段所建 SVM 模型的运算时间;在几种特征波段提取方法中,PCA 和 SPA 所得特征波段数较少,运行时间也较短,而 CARS 提取的特征波段数较多,运行时间也较长,同时准确率也较高(98.25%),基本上可以反映红枣品种的全部信息。

在实际应用中可根据实际需求进行相应的选择。若对速度有一定的要求,CARS 特征提取也可以在满足较高准确率的前提下消耗较短的时间。

5 结 论

本文以新疆金丝大枣、骏枣和滩枣为研究对象,利用高光谱系统获取三种红枣样本的高光谱数据,然后分别采用多种预处理方法(MSC、SNV、1-Der 和 SG 平滑)对红枣样本的高光谱数据进行预处理,并基于多种机器学习算法——LDA、KNN 和 SVM,建立红枣品种鉴别模型,比较了这几种预处理方法对红枣品种鉴别的影响。结果显示:上述几种预处理方法中的 1-Der 预处理方法最优,基于 1-Der 所建立的 LDA、KNN 和 SVM 品种鉴别模型的准确率分别为 76.32%、100%和 100%。接着,本文结合 PCA、SPA 和 CARS 对全波段光谱提取特征波段,所提取的特征波段数分别为 10、13 和 275;然后以全波段和所选取的特征波段作为模型的输入,分别建立基于 LDA、KNN 和 SVM 的红枣品种鉴别模型。比较后发现,基于全波段的 KNN 和 SVM 模型的准确率达到 100%,基于 CARS 特征提取方法所建立的 LDA、KNN 和 SVM 鉴别模型的准确率分别为 85.53%、98.68%和 98.25%。最后,对分别基于全波段和特征波段建立的 SVM 模型的运行时间进行了比较,结果发现,基于全波段的 SVM 模型的运行时间远长于基于特征波段所建 SVM 模型的运行时间。因此,在应用中应根据实际需求,综合考虑准确度、运行时间等因素进行模型的选择。

本文的研究结果表明,基于高光谱技术并结合适当的机器学习建模方法能够无损准确地鉴别新疆

红枣的品种。本文可为进一步开展其他农产品的高光谱分类、品种鉴别与品质检测提供参考。

参 考 文 献

- [1] Goetz A F H, Vane G, Solomon J E, et al. Imaging spectrometry for earth remote sensing[J]. *Science*, 1985, 228(4704): 1147-1153.
 - [2] Tagesson T, Fensholt R, Guiro I, et al. Ecosystem properties of semiarid savanna grassland in West Africa and its relationship with environmental variability[J]. *Global Change Biology*, 2015, 21(1): 250-264.
 - [3] Amodio M L, Capotorto I, Chaudhry M M A, et al. The use of hyperspectral imaging to predict the distribution of internal constituents and to classify edible fennel heads based on the harvest time[J]. *Computers and Electronics in Agriculture*, 2017, 134: 1-10.
 - [4] Li C, Fan P, Jiang K, et al. Melon seed variety identification based on hyperspectral technology combined with discriminant analysis[J]. *Bangladesh Journal of Botany*, 2017, 46(3): 1153-1160.
 - [5] Ravikanth L, Jayas D S, White N D G, et al. Extraction of spectral information from hyperspectral data and application of hyperspectral imaging for food and agricultural products[J]. *Food and Bioprocess Technology*, 2017, 10(1): 1-33.
 - [6] Goto A, Nishikawa J, Kiyotoki S, et al. Use of hyperspectral imaging technology to develop a diagnostic support system for gastric cancer[J]. *Journal of Biomedical Optics*, 2015, 20(1): 016017.
 - [7] Markgraf W, Janssen M W W, Lilienthal J, et al. Hyperspectral imaging for *ex-vivo* organ characterization during normothermic machine perfusion[J]. *European Urology Supplements*, 2018, 17(2): e767.
 - [8] Liu L X, Li M Z, Zhao Z G, et al. Recent advances of hyperspectral imaging application in biomedicine[J]. *Chinese Journal of Lasers*, 2018, 45(2): 0207017.
- 刘立新, 李梦珠, 赵志刚, 等. 高光谱成像技术在生物医学中的应用进展[J]. *中国激光*, 2018, 45(2):

- 0207017.
- [9] Chander S, Gujrati A, Abdul Hakeem K, et al. Water quality assessment of river ganga and chilika lagoon using AVIRIS-NG hyperspectral data [J]. *Current Science*, 2019, 116(7): 1172-1181.
- [10] Yu J W, Cheng Z Q, Zhang J S, et al. An approach to distinguishing between species of trees and crops based on hyperspectral information[J]. *Spectroscopy and Spectral Analysis*, 2018, 38(12): 3890-3896.
虞佳维, 程志庆, 张劲松, 等. 高光谱信息的农林植被种类区分[J]. *光谱学与光谱分析*, 2018, 38(12): 3890-3896.
- [11] Li J B, Tian X, Huang W Q, et al. Application of long-wave near infrared hyperspectral imaging for measurement of soluble solid content (SSC) in pear [J]. *Food Analytical Methods*, 2016, 9(11): 3087-3098.
- [12] Zhang C, Guo C T, Liu F, et al. Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine[J]. *Journal of Food Engineering*, 2016, 179: 11-18.
- [13] Ma T, Li X Z, Inagaki T, et al. Noncontact evaluation of soluble solids content in apples by near-infrared hyperspectral imaging [J]. *Journal of Food Engineering*, 2018, 224: 53-61.
- [14] Rao L B, Pang T, Ji R S, et al. Firmness detection for apples based on hyperspectral imaging technology combined with stack autoencoder-extreme learning machine method [J]. *Laser & Optoelectronics Progress*, 2019, 56(11): 113001.
饶利波, 庞涛, 纪然仕, 等. 基于高光谱成像技术结合堆栈自动编码器-极限学习机方法的苹果硬度检测 [J]. *激光与光电子学进展*, 2019, 56(11): 113001.
- [15] Deng X L, Kong C, Wu W B, et al. Detection of citrus HuangLongBing based on principal component analysis and back propagation neural network [J]. *Acta Photonica Sinica*, 2014, 43(4): 0430002.
邓小玲, 孔晨, 吴伟斌, 等. 基于主成分分析和 BP 神经网络的柑橘黄龙病诊断技术 [J]. *光子学报*, 2014, 43(4): 0430002.
- [16] Sun Y, Gu X Z, Sun K, et al. Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches[J]. *LWT-Food Science and Technology*, 2017, 75: 557-564.
- [17] Fan Y Y, Qiu Z J, Chen J, et al. Identification of varieties of dried red jujubes with near-infrared hyperspectral imaging[J]. *Spectroscopy and Spectral Analysis*, 2017, 37(3): 836-840.
樊阳阳, 裘正军, 陈俭, 等. 基于近红外高光谱成像技术的干制红枣品种鉴别 [J]. *光谱学与光谱分析*, 2017, 37(3): 836-840.
- [18] Pan X Y, Sun L J, Li Y S, et al. Non-destructive classification of apple bruising time based on visible and near-infrared hyperspectral imaging [J]. *Journal of the Science of Food and Agriculture*, 2019, 99(4): 1709-1718.